



Finding the Better of Two Similar Designs by Monte Carlo Techniques

Becker, Peter W.

Published in:
I E E E Transactions on Reliability

Link to article, DOI:
[10.1109/TR.1974.5215264](https://doi.org/10.1109/TR.1974.5215264)

Publication date:
1974

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Becker, P. W. (1974). Finding the Better of Two Similar Designs by Monte Carlo Techniques. *I E E E Transactions on Reliability*, 23(4), 242-246. <https://doi.org/10.1109/TR.1974.5215264>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Finding the Better of Two Similar Designs by Monte Carlo Techniques

PETER W. BECKER, ASSOCIATE MEMBER, IEEE

ABSTRACT—During the iterative design of a system (or circuit) the designer is often faced with the problem of ranking two designs according to some criterion. If the system elements have uncertain values, each system manifestation can be evaluated as to whether or not it meets some performance criterion. The fraction which meets or exceeds the criterion is called the yield. Monte Carlo techniques can be used to simulate the population of systems and thus to estimate the ranking of two designs. The first result presented in the paper is a derivation of the probability that one design is better than another, along with confidence limits for that probability.

If the two designs are topologically the same, i.e., they differ only in the nominal values and actual distributions of true parameter values, then the same set of random numbers can be used for one simulation of each design. Due to the similarity there may be a positive correlation between the 2 results which can then be used to narrow the confidence limits from the crude method mentioned above. The second result is a derivation of these narrow confidence limits.

Reader aids:

Purpose: Widen state of the art

Special math needed for explanations: Elementary statistics

Special math needed for results: None

Results useful to: Designers and reliability engineers

1. INTRODUCTION

1.1 The Problem

The problem of ranking two designs is a recurring one in design optimizations. Though the results presented in this paper have broad applicability, we will restrict the discussion to one important particular problem arising in connection with the optimization of system- and circuit-yield. Yield (strictly speaking, production yield) is the fraction of systems or circuits meeting the stated specifications at the moment they have been built according to the design. During design optimization the designer will again and again be faced with the problem of deciding which of two designs "A" and "B" has the higher yield. During the optimization procedure it is sufficient that the designer be able to *rank* the designs; only towards the end will it become necessary to estimate the yield with accuracy. There are several techniques for estimating yield [1, Chap. 5] and yield-differences, the most powerful being the Monte Carlo methods. This paper shows how two Monte Carlo methods,

"Crude Monte Carlo" and "Correlated Sampling," can be used to rank "A" and "B"; in particular it studies for each method the number of simulations required to achieve a ranking with a prescribed level of confidence.

1.2 Notation

- Y yield (Y is fraction of the population which meets all of the success criteria for the system at hand)
- \hat{Y} an estimate of Y (\hat{Y} is a random variable)
- A, B names of the 2 designs being compared
- a, b subscripts indicating designs A or B
- ΔY $Y_a - Y_b$
- $\Delta \hat{Y}$ an estimate of ΔY ($\Delta \hat{Y}$ is a random variable)
- $\Phi(\cdot)$ the cumulative distribution function, Cdf, of a standard Normal variate
- N the number of simulations of each system
- N_s the number of "successes" in N trials (see Y above)
- x outcome of a subtest on one design or the other, $x = 0$ indicates a failure, $x = 1$ indicates a success (x is a random variable)
- n_{jk} number of times the test outcome is (j,k) ; (j,k) can equal $(0,0)$, $(0,1)$, $(1,0)$ or $(1,1)$
- ρ correlation coefficient
- Σ implies sum from 1 to N
- σ_{ab}^2 variance of $\Delta \hat{Y} \equiv \hat{Y}_a - \hat{Y}_b$ with no correlation possible between any elements of any tests (see Sec. 4)
- σ_{ab}^{*2} variance of $\Delta \hat{Y} \equiv \hat{Y}_a - \hat{Y}_b$ with correlation possible between elements of a pair of tests, and only between elements of single pairs of tests (see Sec. 5)

2. THE MONTE CARLO METHOD

2.1 The Estimator of Yield

With the advent of electronic computers the Monte Carlo method has gained increasing use in simulation problems [2, 3, 4]. In this paper we simulate the process taking place in a production run. The circuits (or systems) are assembled using components selected at random from particular bins. The true values of these components, the input parameters, follow probability distributions of some kind.

Manuscript received June 13, 1972; revised October 2, 1973, and March 19, 1974.

This paper is adapted from one presented at the NATO Conference on "Reliability Testing and Reliability Evaluation," The Hague, Netherlands, September 4-8, 1972 [5].

In translating this situation into an equivalent computer problem, the first step is to write a program for computing the true values of the m output parameters as function of the true values of the n input parameters. Secondly, we obtain (or postulate) the shapes of the probability distributions for the true values of the input parameters. There is no reason to believe that the yield rankings are robust with respect to the pdf's used to describe the distribution of the true values of the input parameters; in other words, use of inaccurate pdf's may well lead to incorrect rankings. Within the computer a random-number generator is used to pick values from these distributions. Once a set of true input parameter values has been obtained, they are inserted in the program and the true values of the circuit output parameters are computed. If any output parameter value violates its constraints the circuit is rejected. The procedure is repeated many times. When a suitable number of circuits, N , has been simulated in this manner, N_s of which met the circuit specification, the yield is estimated by

$$\hat{Y} \equiv N_s/N. \quad (1)$$

Following [1, Sec. 5.5] \hat{Y} is an unbiased estimator of Y . \hat{Y} is a random variable, the values of which are binomially distributed. To simplify the notation let the symbol \hat{Y} indicate both the estimator and the values of the estimator. From [6, Sec. 3.2 and Example 5-35] the mean and the variance of the binomially distributed \hat{Y} are

$$E\{\hat{Y}\} = Y \quad (2)$$

$$\text{Var}\{\hat{Y}\} = Y(1 - Y)/N. \quad (3)$$

The Monte Carlo method

- a) has great intuitive appeal
- b) can handle, without difficulty, any number of output variables, be they linear or nonlinear functions of the input variables
- c) readily takes into account statistical dependence among the input variables.

2.2 Sample Size

The number of trials necessary to obtain some information, in this case about the yield, from a Monte Carlo simulation is important. Naturally the sample size should be small, as otherwise computer run time will be excessive. On the other hand, if the number of trials is too small the statistical confidence level for the computed results is insufficient and the effort is wasted. Certain methods exist for reducing the number of trials; they are based on the exploitation of a feature peculiar to the problem at hand [7, 8]. E.g., in Section 5, we show that a certain covariance may be positive when we rank alternate designs which are similar; later in Section 5 we show how this covariance can be used to reduce the number of trials substantially.

3. THE PROBLEM OF RANKING TWO ALTERNATE DESIGNS

This paper is particularly concerned with the following problem [1]: design a circuit or system which, when mass produced, will meet certain specifications. To arrive at an optimum design an iterative procedure is used which continuously modifies the design, checks to see if the yield has increased, modifies the better design, checks to see if the new modification has increased the yield further, etc. Assume that somewhere during the optimization we must choose one of the two alternate designs, "A" and "B" having the yields Y_a and Y_b . The two designs will often be very similar with regard to topology, nominal values, and tolerances of input parameters, etc. The relevant question is, "is Y_a or Y_b the larger?" In realistic cases Y_a and Y_b can never be determined; only the values of the estimators \hat{Y}_a and \hat{Y}_b can be determined. Then from the sign of $\Delta\hat{Y} \equiv (\hat{Y}_a - \hat{Y}_b)$, which is an unbiased estimator for $(Y_a - Y_b)$, we must decide:

- (i) which is the larger, Y_a or Y_b , and
- (ii) how confident (statistically) we can be that our decisions are correct.

It is sufficient to compute the value of $(\hat{Y}_a - \hat{Y}_b)$ with such accuracy that the sign of the difference is correct (at some pre-determined level of probability), the reason being that it is the proper ranking of the yields that matters during the optimization. When the optimization is terminated it is of course necessary to estimate accurately the yield of the final system.

In the remaining part of the paper we are concerned with the estimator $(\hat{Y}_a - \hat{Y}_b)$. Two different methods: "Crude Monte Carlo," described in Section 4 and "Correlated Sampling" described in Section 5 are used to determine the associated statistical confidence levels. Finally the two methods are compared in Section 6.

4. RANKING BY CRUDE MONTE CARLO

4.1 The Probability Density Function of the Yield Difference

A simple direct approach to generating $\Delta\hat{Y}$ is to generate \hat{Y}_a and \hat{Y}_b separately (see Sec. 2.1) and then to subtract \hat{Y}_b from \hat{Y}_a . Each yield estimate has a binomial distribution [1]. According to de Moivre's theorem [6], for $NY(1 - Y) \gg 1$, a binomial distribution can be approximated by a Normal distribution with the same mean and variance; see (2) and (3). It follows that the variable \hat{Y} is approximately Normally distributed with mean Y and variance $Y(1 - Y)/N$.

Let there be N simulations of each system, all of them mutually statistically independent. Then $\Delta\hat{Y}$ is approximately Normally distributed with mean $[Y_a - Y_b]$ and variance

$$\sigma_{ab}^2 \equiv [Y_a(1 - Y_a) + Y_b(1 - Y_b)]/N.$$

TABLE I

$ \Delta Y $	Y	N
0.1	0.9	70
	0.8	123
0.05	0.9	277
	0.8	492
0.01	0.99	761
	0.9	6915

Minimum sample size for a 5% 2-sided significance test to reject the null hypothesis ($\Delta Y = 0$). It is assumed that $N_a = N_b = N = 2|Y/(\Delta Y)| \cdot |(1 - Y)/(\Delta Y)| \cdot |z_{0.025}|^2$.

Each is easily estimated from the data, but, unfortunately, the distribution of the variance estimate $\hat{\sigma}_{ab}^2$, is not tractable. If it had the chi-square distribution and were independent of the estimate of the mean, one could use the Student's "t" distribution; but those conditions are not satisfied. Nevertheless we will ignore such considerations for the remainder of the paper and assume that N is sufficiently large, so that the random variable

$$z \equiv \frac{\Delta \hat{Y} - \Delta Y}{\hat{\sigma}_{ab}} \quad (4)$$

has a standard Normal distribution. The method of estimating σ_{ab} depends somewhat on the hypothesis made in Secs. 4.2 and 4.3. When (4) applies one can make a significance test of the null hypothesis or one can put confidence limits on $(\Delta \hat{Y} - \Delta Y)$. The two procedures are essentially equivalent since they use the same information.

4.2 Significance Test

For the statistical significance test make the null hypothesis, " $\Delta Y = 0$," and pool the estimates for \hat{Y} to calculate $\hat{\sigma}_{ab}$.

$$\hat{Y} = \frac{N_{sa} + N_{sb}}{2N} \quad (5a)$$

$$\hat{\sigma}_{ab0}^2 = \hat{Y}(1 - \hat{Y})(2/N) \quad (5b)$$

$$\Delta \hat{Y} = \hat{Y}_a - \hat{Y}_b = (N_{sa} - N_{sb})/N \quad (6)$$

Next we calculate z , (4) reduces to (7) under the null hypothesis, namely

$$z = (\Delta \hat{Y})/\hat{\sigma}_{ab0} \quad (7)$$

and we perform the usual significance test. Two-sided tests are more appropriate than single-sided tests, since one

wishes to know which of the two designs is better; and, in advance of the tests, we do not know the sign of $\Delta \hat{Y}$. If $|z|$ is large, the null hypothesis is rejected and ΔY is asserted to have the sign of $\Delta \hat{Y}$. In the case when $\Delta Y \neq 0$ one usually will wish to calculate a confidence interval. Table I shows N versus $|\Delta Y|$ and Y for a 5% 2-sided significance test.

4.3 Confidence Interval for ΔY

Two-sided intervals are often appropriate, but one may wish a single-sided test. The difficulty with the single-sided test is that in advance of the tests, one does not know which side is the right one. To make that decision after the tests is to change the statistical problem originally analyzed. If the null hypothesis is rejected then

$$\hat{\sigma}_{ab}^2 = [\hat{Y}_a(1 - \hat{Y}_a) + \hat{Y}_b(1 - \hat{Y}_b)]/N. \quad (8)$$

$\Delta \hat{Y}$ may be calculated from (6). Recalling the definition of z in (4), confidence statements such as

$$\Delta \hat{Y} + \left(z - \frac{\alpha}{2}\right) \hat{\sigma}_{ab} \leq \Delta Y \leq \Delta \hat{Y} + \left(z + \frac{\alpha}{2}\right) \hat{\sigma}_{ab}$$

are easily made for total confidence α ($0 < \alpha < 1$). If the confidence interval includes zero, one might well wish to assert that $\Delta Y = 0$.

Table I and its associated formula illustrate two noteworthy facts (for $Y > 0.5$), namely

- (i) for fixed $|\Delta Y|$, N decreases as Y increases
- (ii) for fixed Y , N decreases as $|\Delta Y|$ increases

In real life, Y will be fairly high, $|\Delta Y|$ will be fairly small and N will probably be over 1000. The dire need for methods of reducing sample size is as obvious as the methods are limited [7, 8]. Only by exploiting some feature peculiar to the problem at hand can one hope to reduce the sample size. Correlated sampling can reduce sample size when the correlation is positive; Section 5 shows why this is so.

5. RANKING BY CORRELATED MONTE CARLO

Suppose that the two designs are tested in pairs, one implementation of A and one implementation of B . For the test i , we simulate each design with a set of random numbers. The result of test i is a pair of outcomes (x_a, x_b) which can take the values $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$. In this section it is still presumed that the tests are statistically independent, in pairs; i.e., the only possible dependence among the x 's is between those of a pair. In Sec. 4 it was presumed that there was no correlation between x_a and x_b in pair i . That assumption is now dropped. The following relationships are not too difficult to derive

$$E\{\hat{Y}\} = Y$$

$$\sigma^2 \equiv \text{Var}\{\hat{Y}\} = Y(1 - Y)/N$$

$$\begin{aligned}
\text{Cov}\{\hat{Y}_a, \hat{Y}_b\} &= \text{Cov}\left\{\frac{1}{N} \sum x_{ai}, \frac{1}{N} \sum x_{bi}\right\} \\
&= \frac{1}{N^2} \text{Cov}\{\sum x_{ai}, \sum x_{bi}\} = \frac{1}{N^2} \sum \text{Cov}\{x_{ai}, x_{bi}\} \\
&= \frac{1}{N} \text{Cov}\{x_{ai}, x_{bi}\}
\end{aligned}$$

$$\begin{aligned}
\text{Cov}\{x_{ai}, x_{bi}\} &= E\{(x_{ai} - Y_a)(x_{bi} - Y_b)\} \\
&= E\{x_{ai}x_{bi}\} - Y_a Y_b \\
\sigma_{ab}^{*2} &= \text{Var}\{\Delta \hat{Y}\} = \text{Var}\{\hat{Y}_a - \hat{Y}_b\} \\
&= \text{Var}\{\hat{Y}_a\} - 2 \text{Cov}\{\hat{Y}_a, \hat{Y}_b\} + \text{Var}\{\hat{Y}_b\}
\end{aligned}$$

In the following we assume, as in Sec. 4, that the sample size is large enough, so that the true values can be satisfactorily estimated by the sample values (from all $2N$ simulations). Under this assumption we can estimate the terms in the above expression for $\text{Var}\{\Delta \hat{Y}\}$. Also we assume that the central limit theorem holds. Under these assumptions and the null hypothesis

$$z^* \equiv \Delta \hat{Y} / \hat{\sigma}_{ab}^* \text{ and } z \equiv \Delta \hat{Y} / \hat{\sigma}_{ab}$$

both have a standard Normal distribution. The following expressions can be derived using the definitions listed in the first section.

$$\begin{aligned}
\Delta \hat{Y} &= (n_{10} - n_{01})/N \\
\hat{\sigma}_a^2 &= (n_{10} + n_{11})(n_{00} + n_{01})/N^3 \\
\hat{\sigma}_b^2 &= (n_{01} + n_{11})(n_{00} + n_{10})/N^3 \\
\text{Cov}\{x_{ai}, x_{bi}\} &= (n_{11}/N) - (n_{10} + n_{11})(n_{01} + n_{11})/N^2 \\
&= (n_{11}n_{00} - n_{01}n_{10})/N^2 \\
\text{Cov}\{\hat{Y}_a, \hat{Y}_b\} &= (n_{11}n_{00} - n_{01}n_{10})/N^3 \\
\sigma_{ab}^{*2} &= [4n_{10}n_{01} + (n_{11} + n_{00})(n_{01} + n_{10})]/N^3 \quad (9) \\
\sigma_{ab}^2 &= [2(n_{00}n_{11} + n_{10}n_{01}) + (n_{11} + n_{00})(n_{01} + n_{10})]/N^3 \quad (10)
\end{aligned}$$

A moment's reflection reveals that, when $\text{Cov}\{Y_a, Y_b\} = 0$, n_{00}, n_{01}, n_{10} and n_{11} have the following expectations:

$$\begin{aligned}
E\{n_{00}\} &= (1 - Y_a)(1 - Y_b)N, & E\{n_{01}\} &= (1 - Y_a)Y_bN, \\
E\{n_{10}\} &= Y_a(1 - Y_b)N, & E\{n_{11}\} &= Y_a Y_b N.
\end{aligned}$$

Consequently we may expect

$$n_{00}n_{11} \simeq n_{01}n_{10}$$

when separate sets of random numbers were used for all $2N$ design simulations. Inspection of (9) and (10) shows that

$$n_{00}n_{11} = n_{01}n_{10} \leftrightarrow \sigma_{ab}^{*2} = \sigma_{ab}^2.$$

Some further reflection leads to the observation that $\text{Cov}\{x_{ai}, x_{bi}\}$ and $(n_{00}n_{11} - n_{01}n_{10})$ have the same sign. Inspection of (9) and (10) shows that:

$$n_{00}n_{11} > n_{01}n_{10} \leftrightarrow \sigma_{ab}^2 > \sigma_{ab}^{*2}, \leftrightarrow z^* > z.$$

We have thus come to the crux of the matter and the rationale behind "correlated sampling": z^* exceeds z if and only if $\text{Cov}\{x_{ai}, x_{bi}\}$ is positive; the greater the correlation the greater the (z^*/z) -ratio will be. If $z^* > z$, both the significance test and confidence interval mentioned in Sec. 4 will be "sharp"; fewer tests will suffice to discriminate between the two designs. In realistic cases, unfortunately, there is no simple expression, like the one from Table I for the minimum number of tests; but if $\text{Cov}\{x_{ai}, x_{bi}\}$ is known, one can derive an expression [5, (20)].

At this point a word of caution about pseudo-random number generators is in order. By definition a pseudo-random number is generated deterministically; therefore there is some kind of statistical dependence from one number to the next. The reader should ensure that the existing dependence is not of the kind that manifests itself during as few as $2N$ simulated designs; otherwise the results may well be erroneous.

The sample correlation coefficient $\hat{\rho}$ can be checked for being statistically significantly different from zero before using z^* .

$$\begin{aligned}
\hat{\rho} &= \text{Cov}\{\hat{Y}_a, \hat{Y}_b\} / [\hat{\sigma}_a \hat{\sigma}_b] \\
&= [n_{11}n_{00} - n_{01}n_{10}] \\
&\quad \cdot [(n_{11} + n_{10})(n_{11} + n_{01})(n_{00} + n_{10})(n_{00} + n_{01})]^{-1/2}
\end{aligned}$$

$\hat{\rho}\sqrt{N}$ has an approximately standard Normal distribution.

Consequently, a 2-sided test for $\hat{\rho}\sqrt{N}$ being significantly different from zero can be used.

Examples:

	N	n_{11}	n_{00}	n_{10}	n_{01}	\hat{Y}_a	\hat{Y}_b	$\Delta \hat{Y}$	$\hat{\rho}$	$\hat{\rho}\sqrt{N}$	z^*	z
I	200	150	20	20	10	0.85	0.80	0.05	0.5	7	1.8	1.3
II	800	600	80	80	40	0.85	0.80	0.05	0.5	14	3.6	2.6

The correlation coefficient estimate is positive, 0.5 in both I and II. The z^* in I is barely adequate to assert a difference between A and B ; clearly z is even less adequate to assert a difference. The z^* in II is quite adequate to assert a difference and worth-while confidence intervals could be proposed; even without the correlation (i.e. using z instead of z^*) it is reasonable to assert that A and B are different.

6. CONCLUSIONS

Two Monte Carlo methods are feasible for determining the statistical confidence associated with the correct ranking of two system designs, namely, Crude Monte Carlo and

Monte Carlo via Correlated Sampling. When the two design systems, A and B , are similar in topology one can use the same set of random numbers for simulation i of the A -design and simulation i of the B -design. Because of the similarity the outcomes of the two subtests, x_{ai} and x_{bi} will have a positive covariance, $\text{Cov}\{x_{ai}, x_{bi}\} > 0$. However, when the covariance is positive, $\sigma_{ab}^2 > \sigma_{ab}^{*2}$, implying that $z^* > z$. It follows that the method involving correlated sampling is more efficient; it achieves the same statistical confidence with a considerably smaller sample size. Formulas have been derived for determining approximate confidence values for both methods. The aficionado of geometry will appreciate that the same results can be proved geometrically [5].

7. ACKNOWLEDGMENTS

The author thanks Dr. Ralph A. Evans, Mr. Finn Jensen, Dr. Ann F. S. Mitchell and the referees for helpful suggestions.

REFERENCES

- [1] P. W. Becker and F. Jensen, *Design of Systems and Circuits for Maximum Reliability or Maximum Production Yield*; book to be published in 1974 by Polyteknisk Forlag, Copenhagen, Denmark.
- [2] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*, London U.K. Methuen and Co. Ltd., 1967: Section 4.4.
- [3] Yu. A. Shreider, *The Monte Carlo Method*, Long Island City, New York: Pergamon Press, 1966.
- [4] N. P. Buslenko and J. A. Schreider, *Die Monte-Carlo-Methode und ihre Verwirklichung mit elektronischen Digitalrechnern*. Leipzig G.D.R.: B. G. Teubner Verlagsgesellschaft 1964 (translated from Russian).
- [5] P. W. Becker, "Ranking the reliability of two designs by Monte Carlo methods," 1972 NATO Conference Proceedings: Reliability Testing and Reliability Evaluation (The Hague, Netherlands, 4-8 September); Paper no. II-C. Published by E. M. Scheuer, Editor, California State University, Northridge, 1973.
- [6] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, New York: McGraw-Hill, 1965; Section 3.3.
- [7] A. E. Mace, *Sample-Size Determination*, New York: Reinhold, 1964.
- [8] H. Kahn and A. W. Marshall, "Methods of reducing sample size in Monte Carlo computations," *Journal of the Operations Research Society of America*, vol. 1, pp. 263-78; November 1953.

Mailing address:

Dr. Peter W. Becker
Electronics Laboratory
Technical University of Denmark, Bldg. 344
DK-2800 Lyngby, Denmark

Peter W. Becker was born in Copenhagen, Denmark in 1925. He received his M.S. and Dr. Techn. degree in Electrical Engineering from the Technical University of Denmark; presently he is with the Electronics Laboratory of the Technical University of Denmark. Dr. Becker's previous technical experience includes a ten year period with the General Electric Co., Electronics Laboratory, Syracuse, N.Y. He has written several papers in the field of reliability, is interested in pattern recognition, and published in 1968 the book, *Recognition of Patterns*.